Journal of Statistical Planning and Inference 142 (2012) 2762–2778

Contents lists available at SciVerse ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

On the boundary properties of Bernstein polynomial estimators of density and distribution functions

Alexandre Leblanc *

Department of Statistics, University of Manitoba, Winnipeg, Canada R3T 2N2

ARTICLE INFO

Article history:
Received 30 March 2010
Received in revised form
15 March 2012
Accepted 17 March 2012
Available online 27 March 2012

Keywords:
Bernstein polynomials
Density estimation
Mean squared error
Distribution function
Asymptotic properties
Boundary bias

ABSTRACT

For density and distribution functions supported on [0,1], Bernstein polynomial estimators are known to have optimal mean integrated squared error (MISE) properties under the usual smoothness conditions on the function to be estimated. These estimators are also known to be well-behaved in terms of bias: they have uniform bias over the entire unit interval. What is less known, however, is that some of these estimators do experience a boundary effect, but of a different nature than what is seen with the usual kernel estimators.

In this note, we examine the boundary properties of Bernstein estimators of density and distribution functions. Specifically, we show that Bernstein density estimators have decreased bias, but increased variance in the boundary region. In the case of distribution function estimation, we show that Bernstein estimators experience an advantageous boundary effect. Indeed, we prove a particularly impressive property of Bernstein distribution function estimators: they have decreased bias and variance in the boundary region. Finally, we also pay attention to the impact of the so-called *shoulder condition* on the boundary behaviour of these estimators.

1. Introduction

Nonparametric function estimation is undoubtedly an important contribution of Statistics to scientific research. As a testimony to this, a significant amount of literature is available on the subject and many methods have been proposed to solve different problems that fall into this category. One such method that seems to be gaining popularity in recent years, when estimating a function supported on [0,1], is the use of the famous polynomials due to Bernstein (1912) to construct estimators that have advantageous boundary properties. The estimation of densities and distribution functions defined on the unit interval are two important cases where interesting results have been obtained using this approach. Our main purpose here is to further contribute to the understanding of the estimators used in these contexts by studying their boundary behaviour in more detail.

Specifically, assume that a $X_1, X_2, \ldots$ form a sequence of i.i.d. random variables having distribution function F and associated density f supported on the unit interval. Following Babu et al. (2002), the Bernstein estimator of order $m>0$ of F is defined as

$$\hat{F}_{m,n}(x) = \sum_{k=0}^{m} F_n(k/m) P_{k,m}(x), \tag{1}$$

* Tel.: +1 204 474 6273; fax: +1 204 474 7621.
E-mail address: alex_leblanc@umanitoba.ca

0378-3758/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jspi.2012.03.016

where $P_{k,m}(x) = \binom{m}{k} x^k (1-x)^{m-k}$ are binomial probabilities and F_n denotes the empirical distribution function constructed from the first n observations $X_1, X_2, \ldots, X_n$. Throughout this paper, we assume that $m = m_n$ depends on n. The suffix n will, however, be omitted for the sake of simplicity. Note that $\hat{F}_{m,n}$ is a polynomial of degree m with coefficients depending on the data, and thus, leads to very smooth estimates. Note also that $\hat{F}_{m,n}$ yields, with probability one and for any value of m, estimates that are genuine distribution functions on [0,1]. This estimator has been studied by Babu et al. (2002), Babu and Chaubey (2006) and Leblanc (2009, to appear).

Another interesting property of the Bernstein estimator $\hat{F}_{m,n}$ is that, taking its derivative with respect to x, we obtain

$$\frac{d}{dx}\hat{F}_{m,n}(x) = m \sum_{k=0}^{m-1} [F_n([k+1]/m) - F_n(k/m)] P_{k,m-1}(x) = \hat{f}_{m,n}(x), \tag{2}$$

which is the Bernstein density estimator of order m, as it is defined by Babu et al. (2002) and many others. An equivalent form of this estimator was first introduced by Vitale (1975). Note that $\hat{f}_{m,n}$ is a polynomial of degree $m-1$ with coefficients depending on the data. Further note that the density estimator defined above can alternatively be represented as

$$\hat{f}_{m,n}(x) = \sum_{k=0}^{m-1} [F_n([k+1]/m) - F_n(k/m)] \beta_{k+1,m-k}(x), \tag{3}$$

where $\beta_{a,b}(x)$ stands for the Beta density with parameters a and b, that is

$$\beta_{a,b}(x) = \begin{cases} \dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} & \text{for } x \in [0,1], \\ 0 & \text{otherwise}. \end{cases}$$

This is quite interesting as it implies the estimator $\hat{f}_{m,n}$ can be interpreted as a mixture of Beta densities with data driven weights. Another important consequence of this is that, with probability one and for any value of m, $\hat{f}_{m,n}$ yields estimates that are genuine density functions on [0,1].

The density estimator $\hat{f}_{m,n}$ has been studied by Vitale (1975), Babu et al. (2002), Bouezmarni and Rolin (2007) and Leblanc (2010). Generalizations of this estimator, including multivariate and Bayesian extensions, have been considered by Babu and Chaubey (2006), Ghosal (2001), Kakizawa (2004), Kruijer and van der Vaart (2008), Leblanc (2010), Petrone (1999) and Tenbusch (1994).

Other problems of nonparametric function estimation using Bernstein polynomials have also been considered. See, for example, the works of Chang et al. (2005), Choudhuri et al. (2004), Kakizawa (2006), Sancetta (2007) and Tenbusch (1997). We refer the interested reader to Lorentz (1986) for a general introduction to Bernstein polynomials. Finally, we point out that there exists a huge literature on boundary performance and corrections for classical kernel estimators. See, for instance, the work of Jones (1993) and, more recently, Karunamuni and Alberts (2005, 2006). Another interesting reference is the work of Zhang and Karunamuni (2010) who consider the boundary performance of the so-called beta kernel estimators introduced by Chen (1999).

As previously mentioned, our main focus here is to study the boundary behaviour of the Bernstein estimators which we defined in (1) and (2). In other words, we focus on the performance of $\hat{f}_{m,n}$ and $\hat{F}_{m,n}$ at points that are "close" to the boundaries of the unit interval. For this, we let

$$x_{0,m} = \lambda/m \quad \text{and} \quad x_{1,m} = 1 - x_{0,m} = 1 - \lambda/m \tag{4}$$

for some $\lambda \geq 0$, and propose to study the asymptotic behaviour of $\hat{f}_{m,n}(x_{j,m})$ and $\hat{F}_{m,n}(x_{j,m})$ for $j = 0,1$. In the case where $\lambda = 0$, $x_{0,m}$ and $x_{1,m}$ reduce to the boundary points of the unit interval. When $\lambda > 0$, we instead have that $x_{0,m}$ and $x_{1,m}$ are, respectively, points in the left and right boundary regions. Observe that $x_{j,m} \to j$ as $m \to \infty$, for $j = 0,1$.

To better understand what this amounts to in the case where $\lambda > 0$, note that definitions (1) and (2) naturally imply bins of length $1/m$ in which the observed data are grouped (this is why Leblanc, 2010, to appear refers to $1/m$ as the "bandwidth" of Bernstein estimators). Then, by considering $\lambda \in (0,1)$ we are studying the behaviour of the estimators in the first bin (with $x_{0,m}$) and last bin (with $x_{1,m}$). Similarly, by considering $\lambda \in (1,2)$ we are focusing on the second bin (with $x_{0,m}$) and next to last bin (with $x_{1,m}$), and so on. Hence, instead of studying the estimators at a specific x value, we are in essence here looking at the asymptotic performance of the estimators in a specified bin situated in a (shrinking) neighbourhood of the boundaries of the unit interval. The analysis of the boundary properties of the beta kernel estimators done by Chen (1999) and Zhang and Karunamuni (2010) shows some interesting similarities with what we suggest here, but does not offer the same intuitive justification.

Our study will be done in two parts. First, in Section 2, we obtain asymptotic expansions for the bias and variance of the density estimator $\hat{f}_{m,n}$ close to the boundaries of the unit interval. Then, in Section 3, we do the same for the Bernstein estimator $\hat{F}_{m,n}$ of a distribution function.

2. The case of density estimation

We first consider the case of density estimation. For this, we work under the assumption that

$$f \text{ is continuous (and bounded) and admits two continuous and bounded derivatives on } [0,1] \tag{5}$$